DOCUMENT RESUME

ED 079 407 TM 003 011

AUTHOR	Helper, John
TITLE	Materials and Procedures for Assessing Learner Needs in Colorado. Technical Report.
INSTITUTION	Colorado State Dept. of Education, Denver. Div. of Elementary and Secondary Education.
SPONS AGENCY	Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C.
PUB DATE	Apr 73
NOTE	40p.; Draft Copy
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Academic Achievement; *Achievement Tests; Data Analysis; *Educational Needs; Evaluation Methods; *Program Evaluation; *State Programs; Technical Reports; Testing
IDENTIFIERS	*Colorado

ABSTRACT

The Colorado Learner Needs Assessment (CLNA) is described, with operational models for sampling, computer analysis and reporting. The rationale of the program is discussed, followed by its specifications (scope, reporting possibilities, precision and data format), development (meetings, letters, preliminary tryout, sampling and analysis plans, field work plans, administration, scoring and analysis, and mailing results), and uses. (KM)

FILMED FROM BEST AVAILABLE COPY

DRAFT COPY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

ED 079107

EJ 003 011

MATERIALS AND PROCEDURES FOR ASSESSING LEARNER NEEDS IN COLORADO

A Technical Report

. Purposes

. Rationale

. Specifications

. Development

. Uses

Supported in Part by P.L. 89-10 as revised 91-230,
The Elementary and Secondary Education Act, Title III

Prepared by

John Helper, Consultant
Development and Demonstration Services

Colorado Department of Education
Donald D. Woodington, Commissioner
April, 1973

FOREWORD

This report contains technical information regarding materials and procedures for learner needs assessment as developed at the Colorado Department of Education. Full and constructive use of the needs assessment depends a great deal upon the sensitivity and understanding with which it is used. It is to promote such use that this report is intended.

We welcome your comments and questions regarding this information. Through common understandings and cooperative endeavors, we in Colorado may increase our abilities to identify and understand the needs of the young learners in Colorado schools.

Donald D. Woodington,
Commissioner of Education

ACKNOWLEDGEMENTS

Only through extra effort of many persons could these new materials and procedures be developed and made available. Drs. Kenneth D. Hopkins and Gene V Glass and the staff of the Laboratory of Educational Research at the University of Colorado developed the assessment instruments, advised on their administration, supervised the scoring and computer analysis of the results, and provided material for inclusion in this report.

Dr. Barnard D. "Pat" Ryan, Mr. Gene Gullette and other staff of the Boulder Valley Schools assisted in preliminary tryouts of the measures in several different schools in the District.

Also crucial to this project were the several curriculum directors from the Denver area who met several times to review the assessment materials and procedures and to advise on their administration.

Most importantly, persons in the 112 participating school districts--students, teachers, principals, testing coordinators and central office staff--should be acknowledged for their full cooperation in conducting the assessment and reporting their uses of the results.

John W. Helper, Project Coordinator

TABLE OF CONTENTS

	Page
FOREWORD	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
I. PURPOSES	1
II. RATIONALE	1
The Nature of Learner Needs	1
Learner needs defined	2
Individual and group needs	2
Critical needs	2
Learner needs and program needs	2
Readiness	3
Criterion and Norm-referenced Assessment	3
Measurement and Evaluation	4
Measurement	4
Evaluation	5
Multiple Uses	5
State use	5
Local use	5
III. SPECIFICATIONS	6
Scope	6
Reading	7
Language arts	7
Mathematics	7
Attitude toward school	7
Reporting Possibilities	7
Item-level performance	8
Average percent correct	8
District "expected" scores	9
Precision	10
Validity	10
Reliability	11
Data Format	13
Exercise analysis	13
District averages	14
IV. DEVELOPMENT	20
Initial Decision	21
Meetings with District Personnel	21
Contract with University of Colorado	21
Letter of Invitation to District Superintendents	21
Developing Exercise Pools	22
Listing of Participating Districts	22
Preliminary Tryouts	22

Page

Sampling and Analysis Plans..... 23
Final Forms..... 24
Field Work Plans..... 24
Explanatory Materials..... 24
Administration of Instruments..... 24
Scoring and Analysis..... 24
Mailing Results to Districts..... 24
Typical Exercises..... 27

V. USES..... 31
One Example,..... 31
An Overview..... 32
Who used the results?..... 32
Which materials were of most use?..... 33
Were results relevant to local informational needs?..... 34
General Conclusions..... 34

List of Tables

Table — Page

I. Errors of Measurement for District Averages in Skills Assessment 12

II. State Averages, Standard Deviations of District Averages, and Range of District Averages for Each Test (corrected for chance) 13

III. Stanines and Percentiles Corresponding to District Averages 18

IV. Stanines and Percentiles Compared 19

V. Instruction Areas Assessed: Grade 5 25

VI. Instruction Areas Assessed: Grade 11 26

VII. Item-Analysis Data On Basic Skills Instruments - Grade 5 29

VIII. Item-Analysis Data on Basic Skills Instruments - Grade 11 30

List of Figures

Figure		Page
1	Example Exercise Report	8
2	Example District Summary	9
3	Example District Comparison of Obtained and "Expected" Scores	10
4	Sample Printout-District Performance on Skill Exercise	15
5	Sample Printout-District Averages	16
6	Sample Printout-District Expectancies	17
7	The Stanine Scale in Relation to the Normal Distribution, illustrating the Percents of Districts in Various Stanine Categories	19
8	Flow Chart of Major Activities in Developing Colorado's Learner Needs Assessment	20
9	System for Relating Measures of Program Performance to other program Characteristics	31

I. PURPOSES

The Colorado Learner Needs Assessment (CLNA) provides data useful for developing authoritative statements regarding critical needs among learners statewide, in certain pupil population groups, and in those participating school districts where district results are requested. Critical learner needs, defined as "discrepancies between goals and their achievement", guide the development and funding of innovative and exemplary programs, as provided under the Elementary and Secondary Education Act, Title III.

At the same time, the CLNA demonstrates methods for assessment which may be adapted for use in local school districts and other states. Operational models for sampling, computer analysis, and reporting, along with supporting rationale comprise the CLNA, as described in the remainder of this report.

II. RATIONALE

The Colorado Learner Needs Assessment (CLNA) is based on certain assumptions regarding: (1) the nature of learner needs, (2) criterion and norm-referenced assessment, (3) measurement and evaluation, and (4) multiple uses of a single assessment.

The Nature of Learner Needs

Certain commonly-held views on what children need to learn in school are stated in goals such as this one, which is one of several adopted by the Colorado State Board of Education in 1971:

> Each student in the state has the opportunity to acquire THE TECHNIQUES OF LEARNING which make discovery of knowledge and wisdom a functional, exciting and lifelong process.

The CLNA provides a means to determine learner needs in relation to general, long-range educational goals such as the one above. Doing so requires certain assumptions to be made regarding definition of "learner needs", relationships

of individual and group needs, how to identify critical learner needs, a distinction between learner and program needs, and the importance of learner readiness in determining learner needs.

"Learner needs" defined. Central to the CLNA is the definition of "learner needs" as "the discrepancy between goals and their achievement." Basic to this definition is the assumption that particular skills, attitudes and knowledge which enable progress toward long-range goals can be identified by professional judgment, based on what is known about how children learn. Once broad, long-range goals are restated into more specific objectives, appropriate measures can be developed.

Individual and group needs. The CLNA was designed to yield information of the educational needs of groups of learners rather than individual students. This is not intended in any way to minimize the importance of needs of individual students; rather, the intent is to provide general indicators of group needs to guide district program development, not individual student information for planning specific lessons.

Critical needs. Learner needs may be considered to be critical when they seriously jeopardize prospects for achievement of long-range educational goals. Criticality increases with (a) the importance of a particular skill, knowledge or attitude in achieving educational goals; (b) the relative numbers or percentages of students lacking such important skills, knowledge, and attitudes; and (c) the concentration of many needs among certain students. Whether a need is critical calls for a judgment based on cooperative deliberation over results from the CLNA and other relevant data. Such deliberations often result in a better understanding of learner needs by the persons involved.

Learner needs and program needs. A clear distinction between learner needs and program needs is necessary for conducting an accurate assessment of learner needs and for effectively allocating resources to meet these respective

needs. Once learner needs can be accurately and comprehensively assessed, program needs in terms of materials, teacher capabilities and other resources can be specified. The teacher who tells her class, "This test is to help me see how I'm doing as much as it is to help you see how you are doing," illustrates one constructive view of the relationship between learner needs and program needs; she is willing to modify her program on the basis of assessed learner needs.

Readiness. The CLNA was developed according to generally expected or average levels of educational development. The user must consider the appropriateness of the CLNA for a particular group of students before stating conclusively that those students need to learn specific things found to be lacking by the CLNA. A review of the group's past learning experiences and achievements will provide an indication of a general readiness level, which should be considered in developing statements of learner needs.

Criterion and Norm-referenced Assessment

To determine the extent to which certain goals and performance objectives are being achieved, many different types of assessment may be used. The CLNA utilizes two of the more common approaches to assessment--criterion-referenced and norm-referenced

Norm-referenced measures are constructed to distinguish between good spellers and poor spellers, students with large and small vocabularies, with strong and weak mathematical skills, with positive and negative attitudes, and so on. Exercises comprising norm-referenced instruments are selected on the basis of their power to make such distinctions. Instruments made up of such exercises are administered to a sample group (i.e., students in grade 5) and from the obtained score distribution are derived standardized scales, such as stanines and percentiles, by which scores from subsequent testing may be compared.

Criterion-referenced tests, on the other hand, are made up of exercises which indicate in specific detail whether students can or cannot perform certain tasks. An exercise or group of exercises, for example, may be developed to indicate whether or not students can add two three-digit numbers, with carrying. Instead of being chosen for their power to differentiate among skillful and non-skillful students as for norm-referenced tests, exercises for criterion-referenced tests are chosen to elicit skills, knowledge or attitudes as specified in commonly-desired objectives.

Results from both norm and criterion-referenced tests may be interpreted in relation to goals and objectives sought in the school program. Norm-referenced results may be most useful in summarizing learner needs in terms of general goals while criterion-referenced interpretations may be most useful for more detailed diagnosis of learner needs in terms of specific performance objectives. The CLNA was developed to yield both norm and criterion-referenced data.

Measurement and Evaluation

Two different but related operations comprise the needs assessment process of the CLNA--measurement and evaluation. Generally, the Colorado Department of Education provides the CLNA as a means for measurement, and local districts evaluate their own results in terms of local goals and objectives. State-local cooperation is essential to constructive applications of the CLNA.

Measurement. Obtaining reliable and relevant data on learner needs requires specifying the skills, knowledge and attitudes to be measured, selecting appropriate means of measurement, gathering data according to standard procedures, estimating the precision of the data gathered, and setting out the results in a simple and understandable form. Emphasis of the measurement operation is on the technical adequacy of the procedures, adherence thereto, and the estimation of probable measurement error. Precision is sought in measurement activities.

Evaluation. Once the data have been gathered and analyzed, attention turns to their interpretation and constructive uses. This calls for professional judgment regarding (1) the adequacy of the measures in terms of relevance and precision, and (2) the nature and criticality of learner needs indicated in the data. Generally, the rule to follow is this: If important objectives are being measured, if students show weaknesses in demonstrating such learning, and if the measures are of sufficient precision, the user may then conclude that learner needs are indeed present and may proceed to describe these needs in the detail required to guide remedial efforts. Judgment, informed by research and tempered by experience, is sought in evaluation activities.

Multiple Uses

The CLNA was designed to yield data useful at both the state and local district levels of educational responsibility. Specific uses depend on the informational requirements at hand along with the skill and interest of the users to relate the results to commonly-sought goals and performance objectives.

State use. State use of results from the CLNA is in relation to "Educational Goals for Colorado Citizens," providing statewide indicators of learner needs in relation to these commonly-sought goals. Resources made available under the Elementary and Secondary Education Act, Title III have been directed toward developing programs to meet critical needs thus identified.

Local use. The data provided each participating district may also be used for identifying learner needs and planning remediation. Also, procedures of sampling, instrument development and data analysis developed at the Colorado Department of Education may be adapted for use in local school districts, to supplement the data gathered by the CLNA and local testing programs.

III. SPECIFICATIONS

Consistent with the purposes and rationale as described previously, specifications are described here regarding administration, scope, reporting possibilities, precision of the CLNA and data format.

Administration

Specifications regarding administration of the CLNA are as follows:

Student time:	1 hour for skills assessment 20 minutes for attitude assessment No more than 300 students per district per grade
CDE responsibilities:	Offer to each district superintendent Mailing assessment materials Scoring and computer analysis Mailing results to district superintendent Providing suggestions for use of results
District responsibilities:	Requesting to participate Distributing materials to schools Administering instruments Returning materials to CDE Regulating consultant help Reporting irregularities

For districts with more than 200 students per grade level, sampling instructions are provided for each participating school. This entails selecting students via systematic random sampling procedures from a grade-level listing. In the largest district, schools to participate are selected on a stratified random procedure at the Department of Education based on information provided by the district.

Scope

The CLNA provides district-wide indicators of student needs in three common curricular areas and general attitude toward school. There is no intent to assess the extent to which all goals and objectives are being achieved. Rather, the intent is to provide performance information relating to some of the goals and objectives commonly sought across the state.

Reading. Students demonstrate their skills of reading comprehension and vocabulary via multiple-choice exercises which call for skills of:

Recognizing inferences and implications of written material
Distinguishing fact from opinion
Recognizing the purposes of characters portrayed in stories
Identifying principal ideas and topic sentences
Putting sequential events in order
Identifying words most different in meaning from given words

Language arts. The CLNA assesses student skills in four common areas of English language arts:

Spelling--students identify one misspelled word among others spelled correctly
Usage (or grammar)--students identify words and phrases correctly completing a sentence or paragraph or identify one incorrect sentence among two which are grammatically correct
Capitalization--students distinguish among words which are properly capitalized and those which are not
Punctuation--students distinguish between correct and incorrect usage of commas, colons, quotation marks and other commonly used marks of punctuation

Mathematics. Students demonstrate their understanding of arithmatic concepts by selecting correct answers to exercises involving:

Currency, decimals and fractions
Equations
Time, rate and distance
Numerals and number systems
Graphical representation of quantitative relationships

Attitude toward school. A general, district-wide indication of attitude toward school among students at the secondary school level is obtained through a two-page questionnaire including semantic-differential and agree-disagree response modes. Students are asked for their candid opinions (which are to be kept anonymous) regarding various aspects of their school experiences.

Reporting Possibilities

Results, in computer-printout form, for each participating school district are sent to the respective Superintendent. State data and reporting suggestions accompany each district report. Three separate tables of data are provided.

Item-level performance. From percentages of students answering each skill exercise correctly in the district and statewide, such reports as this may be developed locally:

Figure 1.

Example Exercise Report

"On this exercise, 66% of the students in our district at Grade 5 answered this exercise correctly while 85% did so statewide:

The greatest number, using the digits 3, 9, 7, 6 only once is:

a) 3679 b) 9376 c) 7963 *d) 9763 3) 6739

District = 66%
State = 85%

"Based on this and other information already on hand we are considering ways to improve our student's understanding of place value in our elementary mathematics program."

As with all data obtained by sampling, the user needs to know of probable error in the results due to sampling. This is given as a confidence interval within which the "true" district-state discrepancy probably lies. (See Figure 6, page 17)

Average percent correct. The average percent of the exercises in each skill area which were answered correctly summarizes the item-level performance as described previously. Within-district comparisons may be made using the percentile and stanine scales provided by the CLNA as follows:

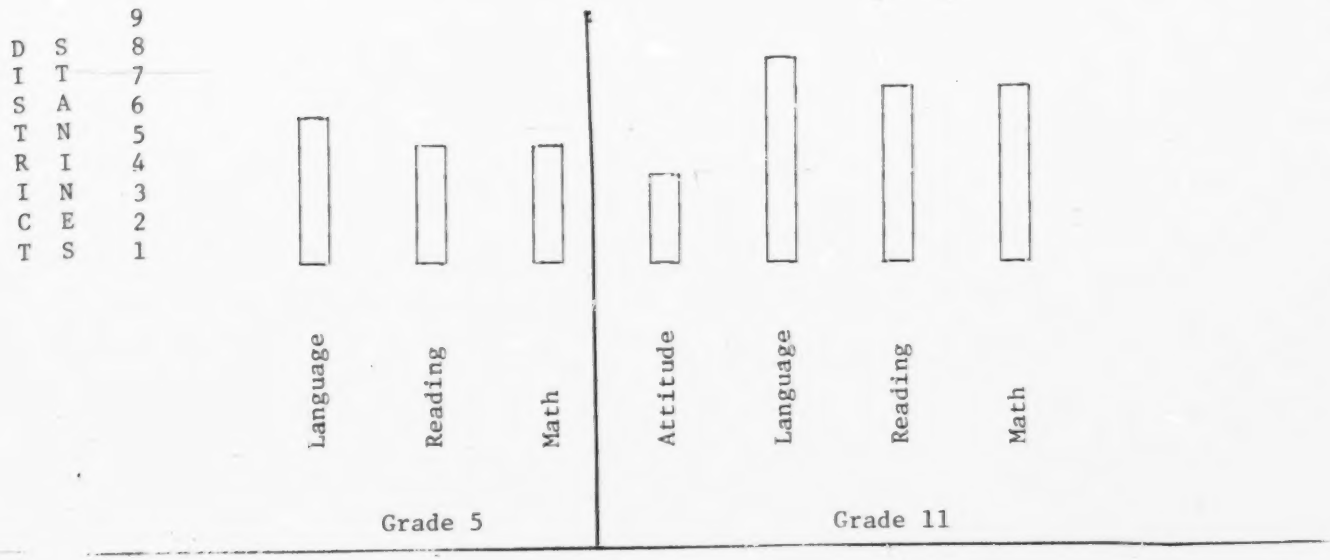
(a) The district percentile, indicating the percentage of districts scoring below the district reported, and

(b) The district stanine which is a normal transformation of the percentile scale, compensating for the tendency of district scores to cluster at the state average.

Within-district comparisons may be shown graphically by such graphs as these taken from the computer printout shown in Figure 5, page 16 of this report.

Figure 2.

Example of a District Summary

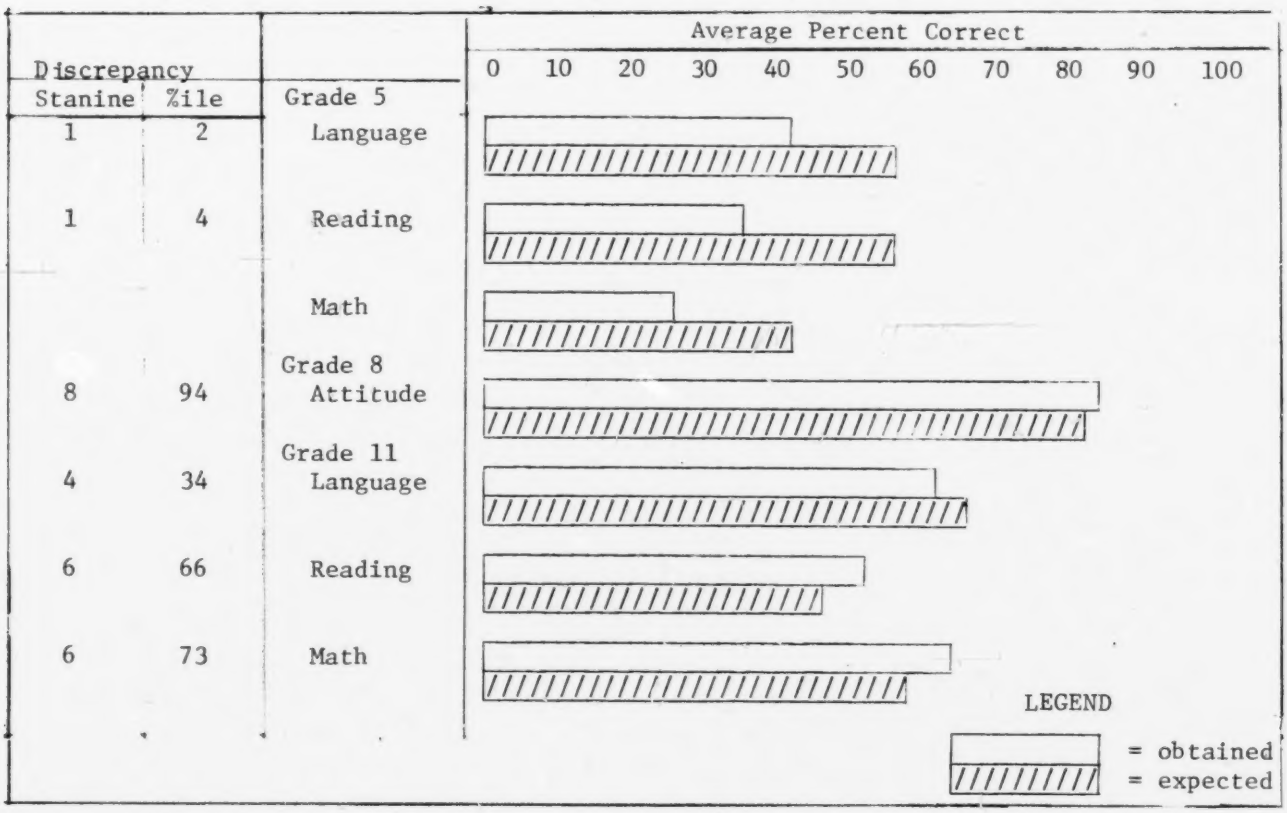


Such comparisons across grade levels among the skill areas may be made with the percentiles and stanines, but not the average percent correct. Relative difficulty of the exercises affects the average percent correct statistic but not percentiles and stanines, which may be considered as common scales permitting within district comparisons. Also, the user should keep in mind that the school attitude score is a scale score, not a "percentage correct", with the higher scale score indicating a more positive attitude toward school.

District "expected" scores. Data are provided each participating district regarding the discrepancy between an "expected" score, based on certain economic factors, and the score actually obtained in a district. Such a graph as that shown on the next page may be developed from this data given in Figure 6, page 16 of this report.

Figure 3.

Example of a District Comparison of Obtained and "Expected" Scores



Precision

Studies of the precision of the CLNA show it to be of more than sufficient validity and reliability to indicate learner needs in each participating district and statewide.

Validity. The CLNA is valid to the extent it taps skills and attitudes commonly sought in a district's educational program. Accordingly, validation procedures started with the development of the instruments, when exercises were discarded, accepted, or revised on the basis of the professional judgments of curriculum experts, teachers, and the developers of the instruments. Further validation of the CLNA was obtained in a questionnaire survey in districts participating in the assessment of 1971.

Judgments of relevance varied from district to district, but indicated a general agreement that the measures covered some (but certainly not all) of the objectives sought in their educational programs.

From the preliminary tryout of the Attitude-Toward-School Inventory, a factor analytic study was conducted to find student response patterns. Eight factors were tentatively identified as contributing to students' general attitudes toward school:

1. Attitude toward school worth.
2. Attitude toward teachers.
3. Attitude toward school relevance.
4. Attitude toward avoidance of school.
5. Undefined.
6. Attitude toward dropping out.
7. Undefined.
8. Attitude toward length of school experience.

Findings from this study are very tentative and, for further validation, require further study of data and experience gained from use of the Inventory statewide. Validity of the Attitude Inventory rests partially on the assumption that the students, unidentified in the scoring of the Inventory, will give candid responses on their attitude toward common aspects of their school experience.

Reliability. District averages obtained on the CLNA are sufficiently reliable that the user can be confident, within certain limits, that the same averages would be obtained on repeated measurements. A single measurement may be considered a sample of a large number of measurements which could be taken and which would probably yield slightly different results. The average of these repeated measurements may be considered to be a "true" score. Information regarding the reliability of the CLNA is set out in Table I.

Table I

Errors of Measurement for District Averages in Skills Assessment

Grade	Test	No. of Districts	Reliability Coefficients	Index of Reliability	Standard error of Measurement
5	Lang. Arts	104	.93	.97	2.5%
5	Reading	104	.95	.98	2.1%
5	Math	104	.90	.95	3.3%
11	Lang. Arts	99	.86	.93	3.9%
11	Reading	99	.86	.93	4.1%
11	Math	99	.92	.96	3.0%

The information in Table I shows that the reliability of the district averages in each of three achievement areas at both grade levels is very high. For example, on grade five Language Arts tests the index of reliability was approximately .97. This indicates that the obtained district means would correlate about .97 with "true" district means, that is, means that would result from perfectly reliable tests--tests containing no random measurement error. The standard error of measurement is the standard deviation of the differences between obtained and "true" district averages. For example, if in district X the "true" average of all examinees on the fifth-grade Language Arts test was 74 percent, the standard error of measurement indicates how close to the mean of 74 percent that the district average would be expected to be. For the grade five Language Arts test, the value of 2.5% indicates that about two-thirds of the districts had averages that did not differ by more than 2.5% from their "true" averages. These small values for the standard errors of measurement show that the obtained district averages differ little from their "true" averages.

From a study conducted by the Laboratory of Educational Research at the University of Colorado the following data were reported regarding the amount

of district-to-district variation around the state average.

Table II.

State Averages, Standard Deviations of District Averages, and Range of District Averages for Each Test (corrected for chance)

Grade		State* Average	Std. Dev. of Dist. Ave.	Range of Dist. Ave.
5	Lang. Arts	57%	8%	36%-75%
5	Reading	55%	9%	27%-76%
5	Math	41%	5%	15%-63%
11	Lang. Arts	65%	7%	47%-84%
11	Reading	47%	5%	13%-68%
11	Math	56%	10%	24%-78%

*State average was defined as the average of the means from the 112 participating Colorado school districts.

The standard deviation values range from 5% to 10% and indicate that there are considerable differences among school districts in the levels of pupil achievement. For example, in Grade 11 math, the standard deviation of 10% indicates that approximately the highest-scoring one-sixth of the districts had averages of 66% or greater, whereas the lowest-scoring districts had averages of 46% or less. Such district-to-district variation permits development of measuring scales which reflect real differences in performance rather than sampling and measurement error.

Data Format

The data are reported to each participating district to enable both criterion and norm-referenced interpretation. Usually a combination of these two approaches is necessary to give a complete and understandable view of learner needs within a district and statewide.

Exercise analysis. As shown in Figure 4, following, statistics for each skill exercise are provided for the participating district as compared with

statewide statistics. To interpret Figure 4, the user of the CLNA needs a copy of the skill exercise to determine what particular skill is assessed. Users should keep in mind that these exercises were chosen both to represent commonly-sought educational objectives and to distinguish between generally skillful and non-skillful learners. For example, a few exercises may indicate capabilities of capitalization among students in the district;a single exercise may indicate student ability to capitalize the first word of a direct quotation.

District averages. Figures 5 and 6 show the printouts for the average percentages of exercises answered correctly in each skill area for the participating district and statewide. On the student Attitude Inventory, a district scale score indicates the extent to which students have generally positive attitudes toward their school environments.

Stanines and percentiles are provided to enable local users of the CLNA to make within-district comparisons among the educational levels and curricular areas represented. Possibilities for reporting learner needs in relation to local goals and objectives were illustrated previously in this report on pages 8 to 10 .

That is, 81% of the district's students who were tested answered item 1 correctly, whereas 85% of the students in the statewide sample answered item 1 correctly. In other words, 4% (81-85) fewer of the district's students answered correctly than for the statewide sample. Considering the sampling error, the "true" percent for the district is probably somewhere between -7% and +1%, that is the discrepancy (-4%) could have resulted for chance as shown by the fact that the confidence includes the value "0", hence the district figure is not reliably different from the state figure.

ITEM LEVEL PERFORMANCE

PERCENT OF PUPILS IN DISTRICT ANSWERING EACH ITEM CORRECTLY.

GRADE 11 LANGUAGE ARTS

See accompanying test copy for specific items

	ITEM	PERCENT CORRECT DIST.	STATE	DISCREPANCY OBTAINED	CONF.INT.
	1	81	85	-4	-7, 1
	2	79	79	0	-3, 5
	3	78	85	-7	-10, -2
SPELLING	4	57	57	0	-4, 5
	5	68	74	-6	-10, -1
	6	50	46	4	-1, 9
	7	92	91	1	-1, 4
	8	63	66	-3	-7, 2
	9	71	68	3	-1, 7
	10	86	93	-7	-9, -3
	11	84	84	0	-3, 4
	12	88	88	0	-2, 4
	13	95	96	-1	-3, 2
	14	71	79	-8	-12, -3
CAPITALIZATION	15	90	92	-2	-5, 1
	16	90	88	2	-1, 5
	17	95	95	0	-2, 2
	18	91	94	-3	-5, 0
	19	94	97	-3	-4, 0
	20	85	89	-4	-6, 0

Figure 4. Sample Printout
District Performance on Skill Exercise

At grade 5, 121 students were tested; at grade 11, 104 students were tested. In the district, the average (mean) score for the Language Arts test was 61%; for the statewide sample, the average score was 57% correct. The district mean on the language arts test exceeded 69% of those for other participating districts. Considering sampling error, the true mean of the district would exceed between 63% and 76% of the other districts. The district performance on this test falls into stanine 6.

DISTRICT PERFORMANCE (AVERAGE PERCENT OF ITEMS CORRECT).

TEST	PUPILS TESTED	AVERAGE PERCENT OF ITEMS CORRECT: DISTRICT	STATE	DISTRICT PERCENTILE: OBTAINED	CONF. INT.	DIST. STANINE
GRADE 5						
LANGUAGE ARTS	121	61	57	69	63- 76	6
READING	121	60	55	70	57- 75	6
MATHEMATICS	121	42	41	50	40- 60	5
GRADE 11						
LANGUAGE ARTS	104	76	65	94	90- 95	8
READING	104	55	47	87	79- 93	7
MATHEMATICS	104	66	56	85	77- 88	7

TEST	PUPILS TESTED	AVERAGE: DISTRICT	STATE	DISTRICT PERCENTILE: OBTAINED	CONF. INT.	DIST. STANINE
GRADE 8						
SCHOOL ATTITUDE	130	77	78	36	29- 47	4

Figure 5. Sample Printout
District Averages

The average score of tested 5th grade students in this district was (41%) on the Language Arts tests, whereas the average score in similar districts was (58%). The discrepancy of (-17%) (58%-41%) was such that only (2%) of the districts had a poorer showing on this test. The confidence interval ((-21% to -13%)) indicates that the discrepancy cannot be attributed to sampling error associated with the sample size. The district's performance at grade 5 was uniformly poor, but at grade 11 the pattern was much better. The student attitudes for grade 8 students were very positive, exceeding (94%) of similar districts.

DISTRICT PERFORMANCE IN RELATION TO EXPECTED PERFORMANCE.

TEST	AVERAGE PERCENT OF ITEMS CORRECT		DISTRICT DISCREPANCY			
	ACTUAL	EXPECTED	OBTAINED	CONF. INT.	PERCENTILE	STANINE
GRADE 5						
LANGUAGE ARTS	(41)	(58)	(-17)	(-21,-13)	(2)	1
READING	38	57	-19	-23,-15	4	1
MATHEMATICS	28	43	-15	-19,-11	3	1
GRADE 11						
LANGUAGE ARTS	64	66	-2	-5, 1	34	4
READING	53	49	4	0, 8	66	6
MATHEMATICS	64	59	5	1, 9	73	6

	AVERAGE		DISTRICT DISCREPANCY			
	ACTUAL	EXPECTED	OBTAINED	CONF. INT.	PERCENTILE	STANINE
GRADE 8						
SCHOOL ATTITUDE	83	78	5	3, 7	(94)	8

Figure 6, Sample Printout
District Expectancies

Table III

Stanines and Percentiles Corresponding to District Averages

Stanine	Percentile	5th Language Arts	5th Reading	5th Math	11th Language Arts	11th Reading	11th Math	Stanine	Percentile	5th Language Arts	5th Reading	5th Math	11th Language Arts	11th Reading	11th Math	Stanine	Percentile	5th Language Arts	5th Reading	5th Math	11th Language Arts	11th Reading	11th Math
4	33	53	51	38	61	44	51																
4	32	53	50	37	61	43	51	6	66	61	59	44	68	50	61	9	99	75	76	63	84	68	78
4	31	53	50	37	61	43	51	6	65	61	59	43	68	49	60	9	98	74	75	61	82	65	77
4	30	52	50	37	61	43	50	6	64	60	59	43	68	49	60	9	97	73	74	61	80	65	76
4	29	52	49	36	61	42	50	6	63	60	58	43	68	49	60	9	96	72	73	60	79	64	75
4	28	52	49	36	60	42	50	6	62	60	58	43	68	49	60	8	95	71	72	58	78	64	74
4	27	52	49	35	60	41	49	6	61	60	58	43	67	49	59	8	94	70	71	58	77	63	73
4	26	51	48	35	60	41	49	6	60	59	58	43	67	49	59	8	93	69	70	57	76	62	73
4	25	51	48	34	60	40	49	5	59	59	58	43	67	49	59	8	92	69	69	56	75	62	72
4	24	51	48	34	59	40	48	5	58	59	57	43	67	49	58	8	91	68	68	54	74	61	71
4	23	51	48	33	59	39	48	5	57	59	57	42	67	48	58	8	90	68	67	54	73	61	71
3	22	50	47	33	59	39	48	5	56	59	57	42	67	48	58	8	89	67	67	54	73	60	71
3	21	50	47	32	59	38	47	5	55	58	57	42	66	48	58	7	88	67	66	53	72	59	70
3	20	50	47	31	58	37	47	5	54	58	56	42	66	48	58	7	87	66	66	52	71	58	70
3	19	50	46	30	58	36	47	5	53	58	56	42	66	48	57	7	86	66	65	52	71	58	69
3	18	49	46	29	58	35	46	5	52	58	56	41	66	47	57	7	85	66	65	51	71	58	68
3	17	49	46	28	58	34	46	5	51	57	55	41	65	47	56	7	84	65	64	51	72	58	66
3	16	49	46	27	56	33	46	5	50	57	55	41	65	47	56	7	83	65	64	51	72	57	66
3	15	48	45	26	56	32	45	5	49	57	55	41	65	47	56	7	82	65	64	50	72	56	65
3	14	48	45	26	56	32	45	5	48	56	54	41	64	47	55	7	81	64	63	49	72	55	65
3	13	47	44	25	55	31	44	5	47	56	54	40	64	46	55	7	80	64	63	49	71	55	65
3	12	47	43	25	55	31	44	5	46	56	54	40	64	46	54	7	79	64	63	48	71	54	64
3	11	46	42	24	54	30	43	5	45	56	53	40	63	46	54	7	78	64	62	48	71	54	64
2	10	46	41	24	54	30	43	5	44	55	53	40	63	46	54	7	77	63	62	47	71	53	64
2	9	45	40	23	53	29	41	5	43	55	53	40	63	46	54	6	76	63	62	47	71	53	63
2	8	45	39	23	52	29	38	5	42	55	53	39	63	45	54	6	75	63	62	46	70	52	63
2	7	44	38	23	51	29	35	5	41	55	52	39	63	45	53	6	74	63	61	46	70	52	63
2	6	43	36	22	51	28	34	5	40	55	52	39	63	45	53	6	73	62	61	46	70	52	62
2	5	42	35	22	50	25	32	4	39	54	52	39	62	45	53	6	72	62	61	45	70	51	62
2	4	40	33	21	49	23	31	4	38	54	52	39	62	45	52	6	71	62	60	45	69	51	62
1	3	39	32	21	49	20	29	4	37	54	51	39	62	45	52	6	70	62	60	45	69	51	62
1	2	38	30	20	48	19	27	4	36	54	51	39	62	45	52	6	69	61	60	45	69	51	61
1	1	37	29	17	48	17	26	4	35	53	51	39	62	45	52	6	68	61	60	44	69	50	61
1	0	36	27	15	47	13	24	4	34	53	51	38	62	44	51	6	67	61	59	44	69	50	61

The stanine scale is a standard score scale that avoids some of the faults of the more easily understood percentile scale. Percentile units near the median are "bunched up," hence the performance difference between the 40th and 50th percentiles is much less than the difference between the 80th and 90th percentiles. The stanine units are equal (except for stanines 1 and 9) hence if stanines are used there is less danger of overinterpreting large differences in percentiles near 50 and underinterpreting small differences in precentiles near the extremes. The figure below illustrates the stanine scale.

Figure 7.

The Stanine Scale in Relation to the Normal Distribution, Illustrating the Percents of Districts in Various Stanine Categories

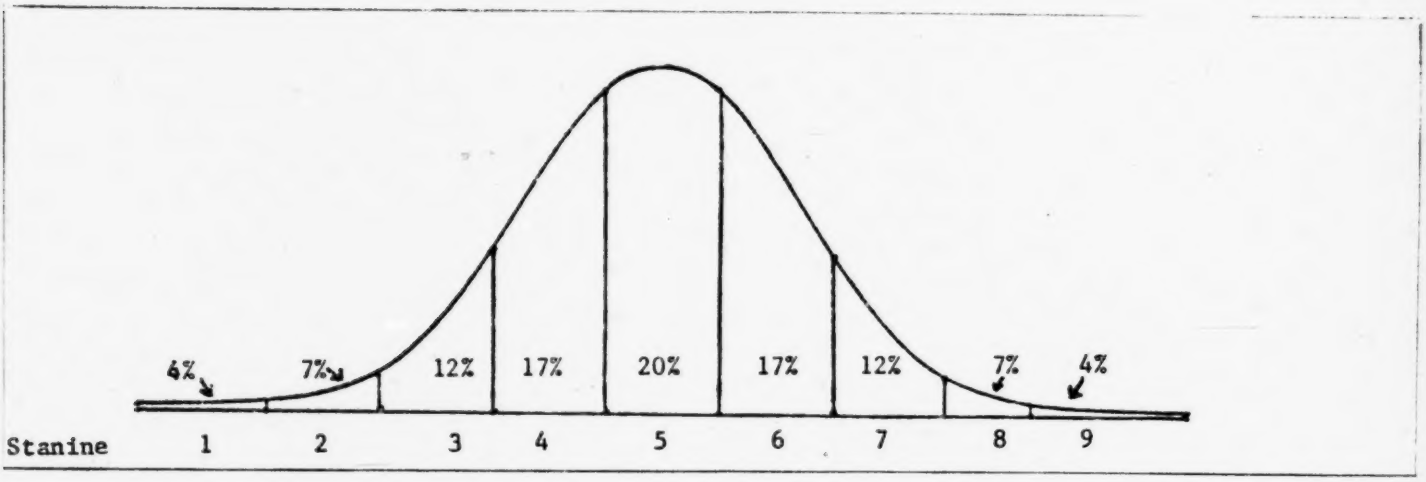


Table IV.

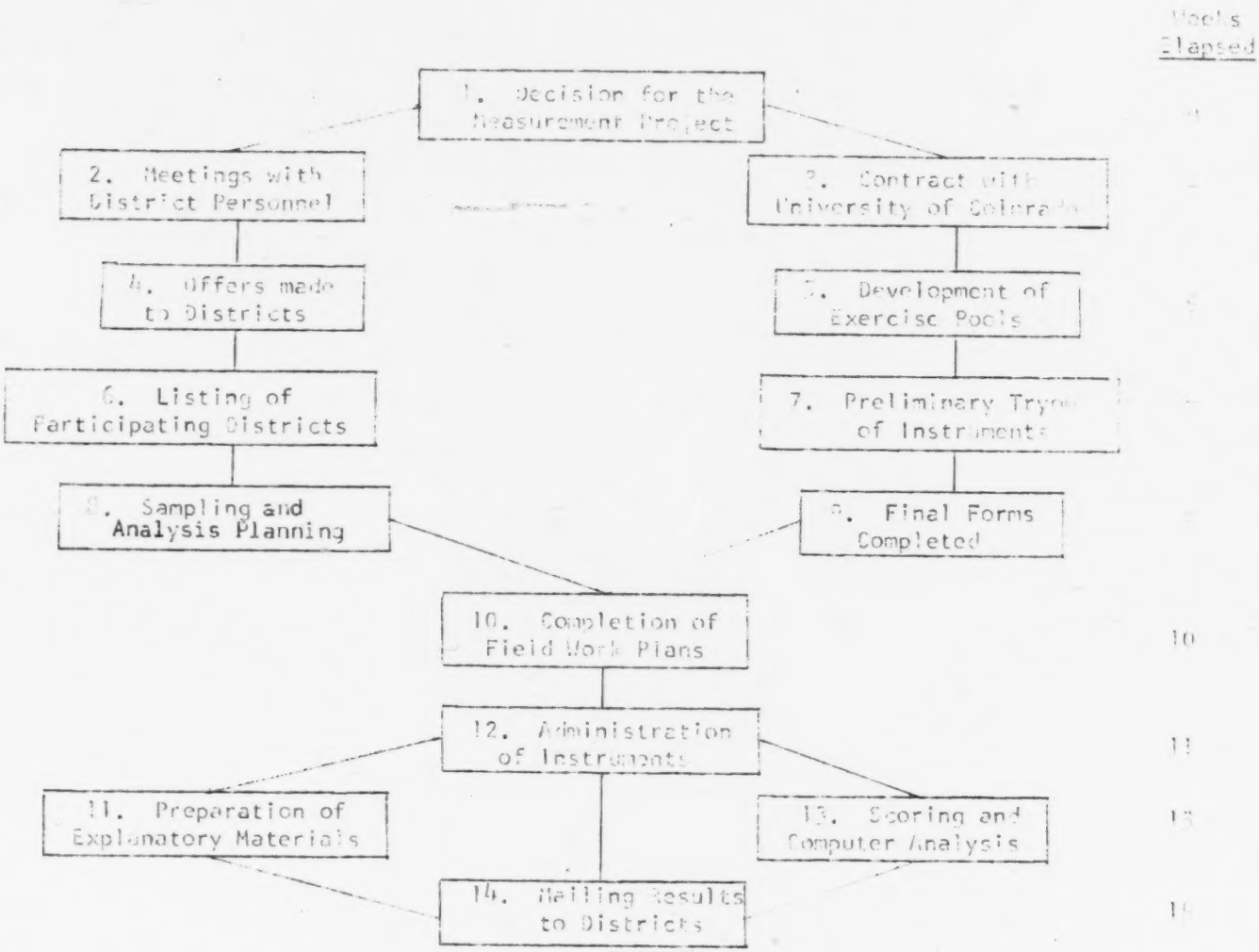
Comparison of Stanines and Percentile Ranges

Stanines	Percentile Ranges
9	Above 95
8	89-95
7	77-88
6	60-76
5	40-59
4	23-39
3	11-22
2	4-10
1	Below 4

IV. DEVELOPMENT

Developing the measures was a cooperative venture, involving personnel from school districts, the Colorado Department of Education, and the Laboratory of Educational Research at the University of Colorado. The major events in the developmental process are indicated in the diagram below and are summarized following.

Figure 8 Flow Chart of Major
Activities in Developing Colorado's Learner Needs Assessment


Weeks Elapsed
1. Decision for the Measurement Project
2. Meetings with District Personnel
3. Contract with University of Colorado
4. Offers made to Districts
5. Development of Exercise Pools
6. Listing of Participating Districts
7. Preliminary Tryout of Instruments
8. Sampling and Analysis Planning
9. Final Forms Completed
10. Completion of Field Work Plans
10
12. Administration of Instruments
11
11. Preparation of Explanatory Materials
13. Scoring and Computer Analysis
13
14. Mailing Results to Districts
14

The Colorado Department of Education had chief responsibilitiy for the following activities: 1, 2, 3, 4, 6, and 14.

Colorado University's Laboratory of Educational Research had chief responsibility for the following activities: 5, 7, 8, 9, 10, 11, and 13.

Personnel in Colorado school districts had chief responsibility for activitiy 12.

1. Initial Decision

The decision to develop new measures was made after an analysis of the data already on hand across the state. The data on hand was found to be too diverse and fragmented to use as common measures of learner needs.

2. Meetings with District Personnel

Concerns expressed in meetings with curriculum and research directors from various districts were that:

1. The program be voluntary
2. Unfair comparisons among districts be avoided
3. The diversity of curricular emphases among districts be considered
4. New data be related to that already on hand.

All of these concerns were observed in the development of the measures.

3. Contract with University of Colorado

The Department entered into a contract with the Laboratory of Educational Research to develop measures with sufficient precision to yield district averages in three pupil skill areas and pupil attitudes toward school. A prior contract was renegotiated to include an exercise-by-exercise analysis of the skill measures.

4. Letter of Invitation to District Superintendents

The Department offered the measures to District Superintendents on a voluntary basis explaining the nature and purpose of the measures.

5. Developing Exercise Pools

The staff of the Laboratory of Educational Research inspected:

- Curricular materials commonly used in Colorado schools
- Objectives judged to be important by Colorado teachers
- Commonly-used achievement tests to develop pools of exercises calling for skills described in these sources

Approximately 120 exercises were developed or collected from the various sources for each program area at each level. At the same time, a preliminary form of the attitude inventory was constructed.

6. Listing of Participating Districts

This was completed approximately 2 weeks after the offer was made to District Superintendents.

7. Preliminary Tryouts

Tryout of exercises in the pools was conducted by the Laboratory of Educational Research in the Boulder, Colorado schools. An effort was made to select schools typifying all those in Colorado in terms of student characteristics. Statistical analyses of the data secured from approximately 125 students for each test included:

Basic skills - (a) item intercorrelations to identify items measuring similar skills

(b) point biserial correlation to identify exercises which generally distinguish between skilled and non-skilled students (see Table VIII)

(c) item difficulty in terms of percentages of students answering correctly.

Attitude inventory - Factor analysis to identify response patterns among students regarding certain aspects of their school experiences.

In addition to these statistical data, judgments were obtained from classroom teachers regarding the relevance of each exercise to common school curricula.

The following topical outline was constructed to guide construction of the final forms.

Grades 5 and 11		Grade 8
Language Arts	Reading	Attitude toward School
Spelling	Vocabulary	Length of school experience
Capitalization	Comprehension	Teachers
Punctuation		Learning
Grammar		School worth
		School relevance
		Avoidance of school

8. Sampling and Analysis Plans

To gain indicators of learner needs for a district, samples were drawn to represent all students reaching certain levels in certain curricular areas as follows:

Educational Level	Curricular Area
Grade 5	Reading, Language Skills, Mathematical Skills
Grade 8	Attitude Toward School
Grade 11	Reading, Language Skills, Mathematical Skills

Selection of students representative of all those served by the program was accomplished by school personnel with instructions provided by the Laboratory of Educational Research. Selection of exercises generally representative of goals and objectives commonly sought in the curricular areas listed above was accomplished by procedures described in step No. 7 as described above.

Concurrently with the sampling plans, the Laboratory of Educational Research planned the computer analysis of results to be consistent with the legal authority and assumptions described in Section II.

9. Final Forms

Based on the procedures described in No. 7 above, exercises were selected to make up internally-consistent instruments to measure the skills and attitudes among students in district programs, according to specifications described in Section III above. Instruments were delivered in camera-ready form by the Laboratory to the Department in time for printing and distribution to districts in December, 1971.

10. Field Work Plans

Field work plans, including test-accounting, monitoring and further communication with participating districts, were worked out upon completion of the sampling plans.

11. Explanatory Materials

Materials explaining the measurement results as well as suggested activities for districts were prepared concurrently with the scoring and computer analysis of the measurement data.

12. Administration of Instruments

School district personnel received materials by mail, administered the instruments according to instructions and returned the completed forms and answer sheets to the Department either by mail or by hand.

13. Scoring and Analysis

Scoring and computer analysis were accomplished at the University of Colorado computer center. Scoring utilized a correction for guessing.

14. Mailing Results to Districts

Each participating Superintendent obtained results for his district before January 1, 1972.

Table V.

Instructional Areas Assessed: Grade 5

Language Arts Elements	Exercise Numbers	No. of Exercises
Spelling	1 through 5	5
Capitalization	6 through 19	14
Punctuation	20 through 26	7
Usage (Grammar)	27 through 39	13
Alphabetizing	40	1
	Total	40

Reading Elements	Exercise Numbers	No. of Exercises
Vocabulary	41 through 57, 72 through 80	26
Comprehension	58 through 71	14
	Total	40

Mathematics Elements	Exercise Numbers	No. of Exercises
Computation	81, 90, 92	3
Concepts	82, 87, 89, 91, 96, 97, 99, 100	8
Application	84, 88, 94, 95, 98	5
Interpreting Graphical Representations	83, 85, 86, 93	4
	Total	20

Grand Total = 100 Exercises

Testing Time = 1 hour (20 minutes per area measured)

Table VI.

Instruction Areas Assessed: Grade 11

Language Arts Elements	Exercise Numbers	No. of Exercises
Spelling	1 through 8	8
Capitalization	9 through 21	13
Punctuation	22 through 26	5
Usage (Grammar)	27 through 40	14
	Total	40

Reading Elements	Exercise Numbers	No. of Exercises
Vocabulary	41 through 53	13
Paragraph Comprehension	54 through 70	17
	Total	30

Mathematics Elements	Exercise Numbers	No. of Exercises
Application	71 through 78, 81 through 90	18
Interpreting Graphical Representation	79, 80	2
	Total	20

Testing Time = 1 hour (20 minutes per area measured)

Grand Total = 90 Exercises

In considering the scope of the CLNA, users must keep in mind that broad program areas and total population of students are represented in the sampling. The results are intended to be general indicators for district-level use and not specific measures for school or classroom use.

Typical Exercises

Grade 11 Math:

Four identical boxes were filled with a total of 144 candy bars. How many were in each box?

a) 24 b) 36 c) 40 d) 576

A driver averaged 60 miles per hour for 3 hours and 15 minutes. How far did he go?

a) 78 miles b) 189 miles c) 195 miles d) 375 miles

The weatherman reported that it was 36° warmer at noon than it had been at sunrise. The noon temperature was 22°. What was the sunrise temperature?

a) -16° b) -14° c) 15° d) 58°

Grade 11 Reading:

Threatened institutions, like endangered species, have often demonstrated remarkable powers of survival. There is the Roman Chursh, the British House of Lords, the German General Staff and the Miss American Pageant. Criticized as lily-white..., demonstrated against by Women's Lib, condescended to by intellectuals and the New York Times (which has been known to spare two paragraphs deep in side to report the winner), Miss American annually blooms like a crop of late summer corn. The second Saturday night in September always finds more than 60 million televiewers tuning in as, live from Atlantic City, Bert Parks opens the last envelope, milks the last drop of suspense, announces the winner and launches the pageant's there song: There She Is.

-Time magazine

Which of the following is not true?

a) The author of the paragraph feels that the New York Times gives generous coverage to the Miss America Pageant.
b) The author feels that the Miss America Pageant is a threatened institution.
c) The author considers the pageant to be over-dramatic.

What does the author mean when he says "Miss America annually blooms like a crop of late summer corn?"

a) The Miss America Pageant is corny.
b) The pageant always take place in late summer.
c) Miss America has most often been selected from "the corn states."

What is meant by the term "lily-white" as used in the paragraph?

a) The pageant is a very clean affair.
b) There are always a lot of white-lilies at the pageant.
c) Few Blacks or other minority persons participate.

Grade 11 Language Arts:

Indicate the word or words which correctly complete the sentence.

_____they went to the game, they did not go to the dance afterwards.

a) While
b) Although

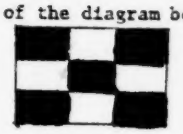
She swam _____, but her golf game was poor.

a) good
b) well

The captain_____hardly see the lighthouse.

Grade 5 Math:

The shaded portions of the diagram below represent what part of the figure?



a) 1/2 b) 5/10 c) 5/9 d) 5/12 e) 4/9

Jane's bike is 1 1/3 yards long. How many feet long is her bike?

a) 4 ft. b) 5 ft. c) 12 ft. d) 13 ft. e) 113 ft.

John bought lunch 17 days in September. Each lunch cost 40 cents, how much did all the lunches cost?

a) $4.80 b) $5.80 c) $6.80 d) $6.90 e) 6.80¢

Grade 5 Language Arts:

Indicate the one word which correctly completes the sentence when it appears in the blank.

The children _____down on the grass.

a) set
b) sat
c) setted

TABLE VII.

Item-Analysis Data
On Basic Skills Instruments

Grade 5

Item No.	p.	r.	Item No.	p.	r.	Item No.	p.	r.
1	.89	.34	35	.51	.38	69	.64	.05
2	.86	.32	36	.61	.42	70	.73	.23
3	.86	.22	37	.74	.49	71	.38	.20
4	.89	.38	38	.75	.46	72	.68	.27
5	.90	.36	39	.58	.53	73	.48	.35
6	.86	.37	40	.52	.41	74	.70	.48
7	.79	.52	41	.85	.37	75	.63	.37
8	.86	.42	42	.79	.33	76	.56	.42
9	.83	.42	43	.48	.35	77	.70	.41
10	.74	.47	44	.61	.29	78	.58	.31
11	.38	.07	45	.80	.29	79	.60	.27
12	.82	.39	46	.86	.34	80	.71	.47
13	.81	.48	47	.63	.20	81	.73	.26
14	.82	.43	48	.72	.35	82	.74	.32
15	.55	.40	49	.79	.39	83	.48	.23
16	.75	.52	50	.77	.32	84	.76	.25
17	.85	.46	51	.58	.21	85	.45	.32
18	.68	.28	52	.54	.32	86	.67	.23
19	.79	.53	53	.64	.40	87	.38	.21
20	.72	.41	54	.77	.35	88	.53	.29
21	.85	.37	55	.76	.38	89	.35	.25
22	.73	.38	56	.43	.14	90	.61	.29
23	.50	.21	57	.49	.30	91	.58	.37
24	.34	.13	58	.78	.35	92	.58	.25
25	.59	.42	59	.56	.24	93	.40	.17
26	.45	.35	60	.56	.27	94	.50	.18
27	.76	.47	61	.88	.34	95	.34	.24
28	.60	.49	62	.77	.22	96	.24	.18
29	.63	.28	63	.79	.21	97	.46	.21
30	.50	.34	64	.77	.32	98	.26	.11
31	.48	.37	65	.86	.35	99	.41	.12
32	.52	.38	66	.80	.36	100	.27	.27
33	.59	.43	67	.67	.31			
34	.77	.46	68	.61	.20			

p = proportion of a sample (N=500) answering item correctly

r = the item-total point-biserial correlation

TABLE VIII.

Item-Analysis Data
On Basic Skills Instruments

Grade 11

Item No.	p.	r.	Item No.	p.	r.	Item No.	p.	r.
1	.70	.65	31	.81	.30	61	.76	.40
2	.68	.67	32	.89	.42	62	.33	.30
3	.86	.23	33	.62	.13	63	.56	.51
4	.51	.52	34	.67	.47	64	.73	.06
5	.61	.63	35	.68	.38	65	.69	.56
6	.38	.36	36	.60	.58	66	.67	.30
7	.80	.61	37	.46	.08	67	.66	.54
8	.55	.57	38	.56	.08	68	.33	.25
9	.73	.09	39	.45	.39	69	.20	.19
10	.83	.56	40	.38	.43	70	.43	.42
11	.81	.44	41	.50	.19	71	.79	.60
12	.77	.69	42	.58	.49	72	.82	.45
13	.84	.69	43	.84	.31	73	.84	.51
14	.81	.14	44	.55	.49	74	.70	.65
15	.87	.52	45	.79	.27	75	.75	.60
16	.88	.34	46	.82	.48	76	.61	.62
17	.81	.73	47	.78	.44	77	.69	.57
18	.90	.40	48	.87	.42	78	.69	.33
19	.82	.68	49	.68	.36	79	.72	.41
20	.77	.57	50	.54	.49	80	.68	.53
21	.80	.21	51	.51	.48	81	.63	.51
22	.73	.16	52	.77	.41	82	.46	.60
23	.69	.58	53	.56	.39	83	.67	.44
24	.70	.53	54	.48	.07	84	.54	.59
25	.83	.39	55	.30	.07	85	.36	.56
26	.76	.71	56	.61	.49	86	.46	.62
27	.46	.49	57	.47	.16	87	.48	.22
28	.67	.21	58	.59	.46	88	.32	.53
29	.76	.40	59	.72	.38	89	.29	.28
30	.91	.42	60	.75	.53	90	.37	.49

p = proportion of a sample (N=500) answering item correctly

r = item-total point-biserial correlation

V. Uses

One Example

Perhaps the most constructive use of the CLNA can be made in school districts where people look at their programs for explanations of why their results are high, low or average. In one school district, for example, the Superintendent found the averages for his district to be considerably below the state averages. His staff told him that perhaps students who could be classified as educationally handicapped had participated and brought the district average down.

The Superintendent's first inclination was to re-compute the district results, excluding the answer sheets of certain students. On further deliberation, however, the Superintendent and his staff decided to look at the district's program for appropriateness for all of the students within the district. Figure 9., following, illustrates this use of the CLNA for program improvement.

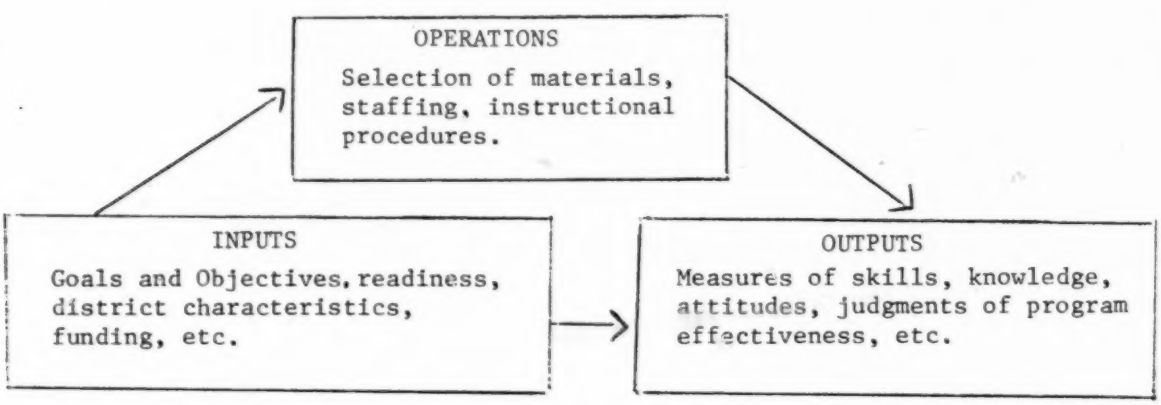

OPERATIONS
Selection of materials, staffing, instructional procedures.
INPUTS
Goals and Objectives, readiness, district characteristics, funding, etc.
OUTPUTS
Measures of skills, knowledge, attitudes, judgments of program effectiveness, etc.

Figure 9.

A System for Relating Measures of

Learner Needs to other program Characteristics*

The district Superintendent and his associates whose deliberations were just described illustrate in principle the model shown in Figure 9. First, they used the results of the measurement as "Output" indicators for their district's program. Second, looking for explanations for the relatively low results, the staff considered "Inputs" in terms of abilities of the students involved. Their final decision, however, was to consider modifications in the "Operations," especially in providing for students who may not be benefiting fully from the present program. (Subsequent measures may reveal gains in the total district program performance due to increased attention to particular needs among certain groups of students.)

Once learner needs can be traced to deficiencies in current educational program practices, then allocation of resources to bolster or modify weak programs is a relatively simple operation. Applications for ESEA, Title III funds, for example, are approved on the basis of whether real learner needs have been identified and whether the proposed program activities are likely to ameliorate these needs.

An Overview

To determine who used what results from the assessment of 1971, and for what purposes, the Department coordinated a questionnaire survey among school personnel who participated in the Fall 1971 assessment. Findings from the survey are given in a special report available from the Department and are summarized below.

Who used the results?

The results appear to be of more use to administrators and teachers than to accountability committees, school boards, and other staff. Administrators in 31 districts made "much use" of the testing results. School

boards in 17 districts made "little use" of the results. The large number (14) of "not applicable" responses for other staff may be due to the absence of staff other than teachers and administrators in many of the smaller participating districts.

Persons from two districts mentioned that results were also useful to counselors and one person mentioned "parents" as a group to whom the results were useful. Three different comments indicated an intent to use the results as time and resources permitted; lack of clerical help in tabulating the responses and the press of other responsibilities were given as reasons for not using the results.

Which materials were of most use?

The average percent correct appeared to be the most useful statistic to local district personnel, followed closely by the percent of students answering each exercise correctly. The confidence interval, perhaps because of its complexity, appeared to be somewhat less useful to district personnel.

The explanatory materials and worksheets received average ratings, indicating considerable usefulness for a majority of districts participating in the voluntary measurement project.

Several respondents requested more detailed analysis, specifically: (1) an item analysis on the attitude survey, (2) results for each subtest (vocabulary, spelling, etc.), (3) school by school analysis when there was more than one school per district administering a given measure, (4) a report of raw scores, and (5) national as well as state comparisons.

Were results relevant to local informational needs?

Perhaps the most important of the questions posed in the survey pertained to the relevance of the measures to local informational needs.

Judgments as to the relevance of the measures to local goals and objectives were distributed all along the "little-much" rating scale, with some tendency toward "much relevance". Approximately 25 percent of the responses (20 of 81) judged "much relevance" of the measures in identifying learner needs.

General Conclusions

1. Because many of the ratings were above average on the utility of the measures, and because at least half of the districts scored below the state average, it would appear that the measures were considered to be useful even in some districts scoring relatively low on the measures.
2. The explanatory materials, worksheets, and instructions to the test coordinators were judged to be useful to a rather high degree in a majority of districts. The actual use of these materials, as indicated in the comments, varied a great deal from considerable use in some districts to more in other districts.
3. Results of the measures appeared to be most useful to administrators and teachers, with administrators using the results more than teachers. This is not surprising as the results were reported for entire districts rather than by school, classroom or individual students.
4. Relevance of the assessment to local goals and performance objectives varied somewhat from district to district, but was generally rated high.
5. Many districts lack the adequate resources to make full and constructive use of the results from the CLNA.